08/349177

HLA-A2.1 BINDING PEPTIDES AND THEIR USES

5   The present application is a continuation in part of USSN 08/159,184 (now abandoned) which is a continuation in part of USSN 08/073,205, (now abandoned) which is a continuation in part of 08/027,146. It is related to USSN 08/205,713. (Pending) All of which are incorporated herein by reference.

10

BACKGROUND OF THE INVENTION

The present invention relates to compositions and methods for preventing, treating or diagnosing a number of pathological states such as viral diseases and cancers. In

15 particular, it provides novel peptides capable of binding selected major histocompatibility complex (MHC) molecules and inducing an immune response.

MHC molecules are classified as either Class I or Class II molecules. Class II MHC molecules are expressed

20 primarily on cells involved in initiating and sustaining immune responses, such as T lymphocytes, B lymphocytes, macrophages, etc. Class II MHC molecules are recognized by helper T lymphocytes and induce proliferation of helper T lymphocytes and amplification of the immune response to the

25 particular immunogenic peptide that is displayed. Class I MHC molecules are expressed on almost all nucleated cells and are recognized by cytotoxic T lymphocytes (CTLs), which then destroy the antigen-bearing cells. CTLs are particularly important in tumor rejection and in fighting viral infections.

30 The CTL recognizes the antigen in the form of a peptide fragment bound to the MHC class I molecules rather than the intact foreign antigen itself. The antigen must normally be endogenously synthesized by the cell, and a portion of the protein antigen is degraded into small peptide

35 fragments in the cytoplasm. Some of these small peptides translocate into a pre-Golgi compartment and interact with class I heavy chains to facilitate proper folding and association with the subunit $\beta2$ microglobulin. The

peptide-MHC class I complex is then routed to the cell surface for expression and potential recognition by specific CTLs.

Investigations of the crystal structure of the human MHC class I molecule, HLA-A2.1, indicate that a peptide

5 binding groove is created by the folding of the α1 and α2 domains of the class I heavy chain (Bjorkman et al., Nature 329:506 (1987). In these investigations, however, the identity of peptides bound to the groove was not determined.

Buus et al., Science 242:1065 (1988) first described

10 a method for acid elution of bound peptides from MHC. Subsequently, Rammensee and his coworkers (Falk et al., Nature 351:290 (1991) have developed an approach to characterize naturally processed peptides bound to class I molecules. Other investigators have successfully achieved direct amino

15 acid sequencing of the more abundant peptides in various HPLC fractions by conventional automated sequencing of peptides eluted from class I molecules of the B type (Jardetzky, et al., Nature 353:326 (1991) and of the A2.1 type by mass spectrometry (Hunt, et al., Science 225:1261 (1992). A review

20 of the characterization of naturally processed peptides in MHC Class I has been presented by Rötzschke and Falk (Rötzschke and Falk, Immunol. Today 12:447 (1991).

Sette et al., Proc. Natl. Acad. Sci. USA 86:3296 (1989) showed that MHC allele specific motifs could be used to

25 predict MHC binding capacity. Schaeffer et al., Proc. Natl. Acad. Sci. USA 86:4649 (1989) showed that MHC binding was related to immunogenicity. Several authors (De Bruijn et al., Eur. J. Immunol., 21:2963-2970 (1991); Pamer et al., 991 Nature 353:852-955 (1991)) have provided preliminary evidence

30 that class I binding motifs can be applied to the identification of potential immunogenic peptides in animal models. Class I motifs specific for a number of human alleles of a given class I isotype have yet to be described. It is desirable that the combined frequencies of these different

35 alleles should be high enough to cover a large fraction or perhaps the majority of the human outbred population.

Despite the developments in the art, the prior art has yet to provide a useful human peptide-based vaccine or

therapeutic agent based on this work.   The present invention provides these and other advantages.

SUMMARY OF THE INVENTION

5   The present invention provides compositions comprising immunogenic peptides having binding motifs for HLA-A2.1 molecules.   The immunogenic peptides, which bind to the appropriate MHC allele, are preferably 9 to 10 residues in length and comprise conserved residues at certain positions

10   such as positions 2 and 9.   Moreover, the peptides do not comprise negative binding residues as defined herein at other positions such as positions 1, 3, 6 and/or 7 in the case of peptides 9 amino acids in length and positions 1, 3, 4, 5, 7, 8 and/or 9 in the case of peptides 10 amino acids in length.

15   The present invention defines positions within a motif enabling the selection of peptides which will bind efficiently to HLA A2.1.

Epitopes on a number of immunogenic target proteins can be identified using the peptides of the invention.

20   Examples of suitable antigens include prostate cancer specific antigen (PSA), hepatitis B core and surface antigens (HBVc, HBVs) hepatitis C antigens, Epstein-Barr virus antigens, human immunodeficiency type-1 virus (HIV1) and papilloma virus antigens.   The peptides are thus useful in pharmaceutical

25   compositions for both _in vivo_ and _ex vivo_ therapeutic and diagnostic applications.

Definitions

The term "peptide" is used interchangeably with

30   "oligopeptide" in the present specification to designate a series of residues, typically L-amino acids, connected one to the other typically by peptide bonds between the alpha-amino and carbonyl groups of adjacent amino acids.   The oligopeptides of the invention are less than about 15 residues

35   in length and usually consist of between about 8 and about 11 residues, preferably 9 or 10 residues.

An "immunogenic peptide" is a peptide which comprises an allele-specific motif such that the peptide will

bind an MHC molecule and induce a CTL response. Immunogenic peptides of the invention are capable of binding to an appropriate HLA-A2.1 molecul and inducing a cytotoxic T cell response against the antigen from which the immunogenic

5   . peptide is derived.

Immunogenic peptides are conveniently identified using the algorithms of the invention. The algorithms are mathematical procedures that produce a score which enables the selection of immunogenic peptides. Typically one uses the

10   algorithmic score with a "binding threshold" to enable selection of peptides that have a high probability of binding at a certain affinity and will in turn be immunogenic. The algorithm is based upon either the effects on MHC binding of a particular amino acid at a particular position of a peptide or

15   the effects on binding of a particular substitution in a motif containing peptide.

A "conserved residue" is an amino acid which occurs in a significantly higher frequency than would be expected by random distribution at a particular position in a peptide.

20   Typically a conserved residue is one where the MHC structure may provide a contact point with the immunogenic peptide. One to three, preferably two, conserved residues within a peptide of defined length defines a motif for an immunogenic peptide. These residues are typically in close contact with the peptide

25   binding groove, with their side chains buried in specific pockets of the groove itself. Typically, an immunogenic peptide will comprise up to three conserved residues, more usually two conserved residues.

As used herein, "negative binding residues" are

30   amino acids which if present at certain positions (for example, positions 1, 3 and/or 7 of a 9-mer) will result in a peptide being a nonbinder or poor binder and in turn fail to be immunogenic i.e. induce a CTL response.

The term "motif" refers to the pattern of residues

35   in a peptide of defined length, usually about 8 to about 11 amino acids, which is recognized by a particular MHC allele. The peptide motifs are typically different for each human MHC

allele and differ in the pattern of the highly conserved residues and negative residues.

The binding motif for an allele can be defined with increasing degrees of precision. In one case, all of the
5 conserved residues are present in the correct positions in a peptide and there are no negative residues in positions 1,3 and/or 7.

The phrases "isolated" or "biologically pure" refer to material which is substantially or essentially free from
10 components which normally accompany it as found in its native state. Thus, the peptides of this invention do not contain materials normally associated with their in situ environment, e.g., MHC I molecules on antigen presenting cells. Even where a protein has been isolated to a homogenous or dominant band,
15 there are trace contaminants in the range of 5-10% of native protein which co-purify with the desired protein. Isolated peptides of this invention do not contain such endogenous co-purified protein.

The term "residue" refers to an amino acid or amino
20 acid mimetic incorporated in an oligopeptide by an amide bond or amide bond mimetic.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention relates to the determination
25 of allele-specific peptide motifs for human Class I MHC (sometimes referred to as HLA) allele subtypes, in particular, peptide motifs recognized by HLA-A2.1 alleles. These motifs are then used to define T cell epitopes from any desired antigen, particularly those associated with human viral
30 diseases, cancers or autoiummune diseases, for which the amino acid sequence of the potential antigen or autoantigen targets is known.

Epitopes on a number of potential target proteins can be identified in this manner. Examples of suitable
35 antigens include prostate specific antigen (PSA), hepatitis B core and surface antigens (HBVc, HBVs) hepatitis C antigens, Epstein-Barr virus antigens, melanoma antigens (e.g., MAGE-1),

human immunodeficiency virus (HIV) antigens and human papilloma virus (HPV) antigens.

Autoimmune associated disorders for which the peptides of the invention may be employed to relieve the symptoms of, treat or prevent the occurrence or reoccurrence of include, for example, multiple sclerosis (MS), rheumatoid arthritis (RA), Sjogren syndrome, scleroderma, polymyositis, dermatomyositis, systemic lupus erythematosus, juvenile rheumatoid arthritis, ankylosing spondylitis, myasthenia gravis (MG), bullous pemphigoid (antibodies to basement membrane at dermal-epidermal junction), pemphigus (antibodies to mucopolysaccharide protein complex or intracellular cement substance), glomerulonephritis (antibodies to glomerular basement membrane), Goodpasture's syndrome, autoimmune hemolytic anemia (antibodies to erythrocytes), Hashimoto's disease (antibodies to thyroid), pernicious anemia (antibodies to intrinsic factor), idiopathic thrombocytopenic purpura (antibodies to platelets), Grave's disease, and Addison's disease (antibodies to thyroglobulin), and the like.

The autoantigens associated with a number of these diseases have been identified. For example, in experimentally induced autoimmune diseases, antigens involved in pathogenesis have been characterized: in arthritis in rat and mouse, native type-II collagen is identified in collagen-induced arthritis, and mycobacterial heat shock protein in adjuvant arthritis; thyroglobulin has been identified in experimental allergic thyroiditis (EAT) in mouse; acetyl choline receptor (AChR) in experimental allergic myasthenia gravis (EAMG); and myelin basic protein (MBP) and proteolipid protein (PLP) in experimental allergic encephalomyelitis (EAE) in mouse and rat. In addition, target antigens have been identified in humans: type-II collagen in human rheumatoid arthritis; and acetyl choline receptor in myasthenia gravis.

Peptides comprising the epitopes from these antigens are synthesized and then tested for their ability to bind to the appropriate MHC molecules in assays using, for example, purified class I molecules and radioiodonated peptides and/or cells expressing empty class I molecules by, for instance,

immunofluorescent staining and flow microfluorometry, peptide-dependent class I assembly assays, and inhibition of CTL recognition by peptide competition. Those peptides that bind to the class I molecule are further evaluated for their

5   ability to serve as targets for CTLs derived from infected or immunized individuals, as well as for their capacity to induce primary *in* *vitro* or *in* *vivo* CTL responses that can give rise to CTL populations capable of reacting with virally infected target cells or tumor cells as potential therapeutic agents.

10   The MHC class I antigens are encoded by the HLA-A, B, and C loci. HLA-A and B antigens are expressed at the cell surface at approximately equal densities, whereas the expression of HLA-C is significantly lower (perhaps as much as 10-fold lower). Each of these loci have a number of alleles.

15   The peptide binding motifs of the invention are relatively specific for each allelic subtype.

          For peptide-based vaccines, the peptides of the present invention preferably comprise a motif recognized by an MHC I molecule having a wide distribution in the human

20   population. Since the MHC alleles occur at different frequencies within different ethnic groups and races, the choice of target MHC allele may depend upon the target population. Table 1 shows the frequency of various alleles at the HLA-A locus products among different races. For instance,

25   the majority of the Caucasoid population can be covered by peptides which bind to four HLA-A allele subtypes, specifically HLA-A2.1, A1, A3.2, and A24.1. Similarly, the majority of the Asian population is encompassed with the addition of peptides binding to a fifth allele HLA-A11.2.

30

TABLE 1

A Allele/Subtype	N(69)[*]	A(54)	C(502)
A1	10.1(7)	1.8(1)	27.4(138)
A2.1	11.5(8)	37.0(20)	39.8(199)
A2.2	10.1(7)	0	3.3(17)
A2.3	1.4(1)	5.5(3)	0.8(4)
A2.4	–	–	–
A2.5	–	–	–
A3.1	1.4(1)	0	0.2(0)
A3.2	5.7(4)	5.5(3)	21.5(108)
A11.1	0	5.5(3)	0
A11.2	5.7(4)	31.4(17)	8.7(44)
A11.3	0	3.7(2)	0
A23	4.3(3)	–	3.9(20)
A24	2.9(2)	27.7(15)	15.3(77)
A24.2	–	–	–
A24.3	–	–	–
A25	1.4(1)	–	6.9(35)
A26.1	4.3(3)	9.2(5)	5.9(30)
A26.2	7.2(5)	–	1.0(5)
A26V	–	3.7(2)	–
A28.1	10.1(7)	–	1.6(8)
A28.2	1.4(1)	–	7.5(38)
A29.1	1.4(1)	–	1.4(7)
A29.2	10.1(7)	1.8(1)	5.3(27)
A30.1	8.6(6)	–	4.9(25)
A30.2	1.4(1)	–	0.2(1)
A30.3	7.2(5)	–	3.9(20)
A31	4.3(3)	7.4(4)	6.9(35)
A32	2.8(2)	–	7.1(36)
Aw33.1	8.6(6)	–	2.5(13)
Aw33.2	2.8(2)	16.6(9)	1.2(6)
Aw34.1	1.4(1)	–	–
Aw34.2	14.5(10)	–	0.8(4)
Aw36	5.9(4)	–	–

Table compiled from B. DuPont, Immunobiology of HLA, Vol. I, Histocompatibility Testing 1987, Springer-Verlag, New York 1989.

[*] N – negroid; A = Asian; C = caucasoid. Numbers in parenthesis represent the number of individuals included in the analysis.

The nomenclature used to describe peptide compounds follows the conventional practice wherein the amino group is presented to the left (the N-terminus) and the carboxyl group to the right (the C-terminus) of each amino acid residue. In the formulae representing selected specific embodiments of the present invention, the amino- and carboxyl-terminal groups, although not specifically shown, are in the form they would

assume at physiologic pH values, unless otherwise specified. In the amino acid structure formulae, each residue is generally represented by standard three letter or single letter designations. The L-form of an amino acid residue is

5   represented by a capital single letter or a capital first letter of a three-letter symbol, and the D-form for those amino acids having D-forms is represented by a lower case single letter or a lower case three letter symbol. Glycine has no asymmetric carbon atom and is simply referred to as

10   "Gly" or G.

The procedures used to identify peptides of the present invention generally follow the methods disclosed in Falk et al., <u>Nature</u> 351:290 (1991), which is incorporated herein by reference. Briefly, the methods involve large-scale

15   isolation of MHC class I molecules, typically by immunoprecipitation or affinity chromatography, from the appropriate cell or cell line. Examples of other methods for isolation of the desired MHC molecule equally well known to the artisan include ion exchange chromatography, lectin

20   chromatography, size exclusion, high performance ligand chromatography, and a combination of all of the above techniques.

In the typical case, immunoprecipitation is used to isolate the desired allele. A number of protocols can be

25   used, depending upon the specificity of the antibodies used. For example, allele-specific mAb reagents can be used for the affinity purification of the HLA-A, HLA-B$_1$, and HLA-C molecules. Several mAb reagents for the isolation of HLA-A molecules are available. The monoclonal BB7.2 is suitable for

30   isolating HLA-A2 molecules. Affinity columns prepared with these mAbs using standard techniques are successfully used to purify the respective HLA-A allele products.

In addition to allele-specific mAbs, broadly reactive anti-HLA-A, B, C mAbs, such as W6/32 and B9.12.1, and

35   one anti-HLA-B, C mAb, B1.23.2, could be used in alternative affinity purification protocols as described in the example section below.   The peptides bound to the peptide binding groove of the isolated MHC molecules are eluted

typically using acid treatment. Peptides can also be dissociated from class I molecules by a variety of standard denaturing means, such as heat, pH, detergents, salts, chaotropic agents, or a combination thereof.

5   Peptide fractions are further separated from the MHC molecules by reversed-phase high performance liquid chromatography (HPLC) and sequenced. Peptides can be separated by a variety of other standard means well known to the artisan, including filtration, ultrafiltration,

10   electrophoresis, size chromatography, precipitation with specific antibodies, ion exchange chromatography, isoelectrofocusing, and the like.

Sequencing of the isolated peptides can be performed according to standard techniques such as Edman degradation

15   (Hunkapiller, M.W., et al., Methods Enzymol. 91, 399 [1983]). Other methods suitable for sequencing include mass spectrometry sequencing of individual peptides as previously described (Hunt, et al., Science 225:1261 (1992), which is incorporated herein by reference). Amino acid sequencing of

20   bulk heterogenous peptides (e.g., pooled HPLC fractions) from different class I molecules typically reveals a characteristic sequence motif for each class I allele.

Definition of motifs specific for different class I alleles allows the identification of potential peptide

25   epitopes from an antigenic protein whose amino acid sequence is known. Typically, identification of potential peptide epitopes is initially carried out using a computer to scan the amino acid sequence of a desired antigen for the presence of motifs. The epitopic sequences are then synthesized. The

30   capacity to bind MHC Class molecules is measured in a variety of different ways. One means is a Class I molecule binding assay as described in Example 8, below. Other alternatives described in the literature include inhibition of antigen presentation (Sette, et al., J. Immunol. 141:3893 (1991), in

35   vitro assembly assays (Townsend, et al., Cell 62:285 (1990), and FACS based assays using mutated ells, such as RMA.S (Melief, et al., Eur. J. Immunol. 21:2963 (1991)).

Next, peptides that test positive in the MHC class I binding assay are assayed for the ability of the peptides to induce specific CTL responses _in vitro_. For instance, Antigen-presenting cells that have been incubated with a

5 peptide can be assayed for the ability to induce CTL responses in responder cell populations. Antigen-presenting cells can be normal cells such as peripheral blood mononuclear cells or dendritic cells (Inaba, et al., _J. Exp. Med._ 166:182 (1987); Boog, _Eur. J. Immunol_. 18:219 [1988]).

10 Alternatively, mutant mammalian cell lines that are deficient in their ability to load class I molecules with internally processed peptides, such as the mouse cell lines RMA-S (Kärre, et al.. _Nature_, 319:675 (1986); Ljunggren, et al., _Eur. J. Immunol_. 21:2963-2970 (1991)), and the human

15 somatic T cell hybrid, T-2 (Cerundolo, et al., _Nature_ 345:449-452 (1990)) and which have been transfected with the appropriate human class I genes are conveniently used, when peptide is added to them, to test for the capacity of the peptide to induce _in vitro_ primary CTL responses. Other

20 eukaryotic cell lines which could be used include various insect cell lines such as mosquito larvae (ATCC cell lines CCL 125, 126, 1660, 1591, 6585, 6586), silkworm (ATTC CRL 8851), armyworm (ATCC CRL 1711), moth (ATCC CCL 80) and Drosophila cell lines such as a Schneider cell line (see Schneider _J._

25 _Embryol. Exp. Morphol._ 27:353-365 [1927]).

Peripheral blood lymphocytes are conveniently isolated following simple venipuncture or leukapheresis of normal donors or patients and used as the responder cell sources of CTL precursors. In one embodiment, the appropriate

30 antigen-presenting cells are incubated with 10-100 μM of peptide in serum-free media for 4 hours under appropriate culture conditions. The peptide-loaded antigen-presenting cells are then incubated with the responder cell populations _in vitro_ for 7 to 10 days under optimized culture conditions.

35 Positive CTL activation can be determined by assaying the cultures for the presence of CTLs that kill radiolabeled target cells, both specific peptide-pulsed targets as well as target cells expressing endogenously processed form of the

relevant virus or tumor antigen from which the peptide sequence was derived.

Specificity and MHC restriction of the CTL is determined by testing against different peptide target cells
5   expressing appropriate or inappropriate human MHC class I. The peptides that test positive in the MHC binding assays and give rise to specific CTL responses are referred to herein as immunogenic peptides.

The immunogenic peptides can be prepared
10   synthetically, or by recombinant DNA technology or from natural sources such as whole viruses or tumors. Although the peptide will preferably be substantially free of other naturally occurring host cell proteins and fragments thereof, in some embodiments the peptides can be synthetically
15   conjugated to native fragments or particles.

The polypeptides or peptides can be a variety of lengths, either in their neutral (uncharged) forms or in forms which are salts, and either free of modifications such as glycosylation, side chain oxidation, or phosphorylation or
20   containing these modifications, subject to the condition that the modification not destroy the biological activity of the polypeptides as herein described.

Desirably, the peptide will be as small as possible while still maintaining substantially all of the biological
25   activity of the large peptide. When possible, it may be desirable to optimize peptides of the invention to a length of 9 or 10 amino acid residues, commensurate in size with endogenously processed viral peptides or tumor cell peptides that are bound to MHC class I molecules on the cell surface.
30   Peptides having the desired activity may be modified as necessary to provide certain desired attributes, e.g., improved pharmacological characteristics, while increasing or at least retaining substantially all of the biological activity of the unmodified peptide to bind the desired MHC
35   molecule and activate the appropriate T cell. For instance, the peptides may be subject to various changes, such as substitutions, either conservative or non-conservative, where such changes might provide for certain advantages in their

use, such as improved MHC binding. By conservative
substitutions is meant replacing an amino acid residue with
another which is biologically and/or chemically similar, e.g.,
one hydrophobic residue for another, or one polar residue for

5   another. The substitutions include combinations such as Gly,
Ala; Val, Ile, Leu, Met; Asp, Glu; Asn, Gln; Ser, Thr; Lys,
Arg; and Phe, Tyr. The effect of single amino acid
substitutions may also be probed using D-amino acids. Such
modifications may be made using well known peptide synthesis

10   procedures, as described in e.g., Merrifield, _Science_ 232:341-
347 (1986), Barany and Merrifield, _The Peptides_, Gross and
Meienhofer, eds. (N.Y., Academic Press), pp. 1-284 (1979); and
Stewart and Young, _Solid Phase Peptide Synthesis_, (Rockford,
Ill., Pierce), 2d Ed. (1984), incorporated by reference

15   herein.

        The peptides can also be modified by extending or
decreasing the compound's amino acid sequence, e.g., by the
addition or deletion of amino acids. The peptides or analogs
of the invention can also be modified by altering the order or

20   composition of certain residues, it being readily appreciated
that certain amino acid residues essential for biological
activity, e.g., those at critical contact sites or conserved
residues, may generally not be altered without an adverse
effect on biological activity. The non-critical amino acids

25   need not be limited to those naturally occurring in proteins,
such as L-α-amino acids, or their D-isomers, but may include
non-natural amino acids as well, such as β-γ-δ-amino acids, as
well as many derivatives of L-α-amino acids.

        Typically, a series of peptides with single amino

30   acid substitutions are employed to determine the effect of
electrostatic charge, hydrophobicity, etc. on binding. For
instance, a series of positively charged (e.g., Lys or Arg) or
negatively charged (e.g., Glu) amino acid substitutions are
made along the length of the peptide revealing different

35   patterns of sensitivity towards various MHC molecules and T
cell receptors. In addition, multiple substitutions using
small, relatively neutral moieties such as Ala, Gly, Pro, or
similar residues may be employed. The substitutions may be

homo-oligomers or hetero-oligomers. The number and types of residues which are substituted or added depend on the spacing necessary between essential contact points and certain functional attributes which are sought (e.g., hydrophobicity

5   versus hydrophilicity). Increased binding affinity for an MHC molecule or T cell receptor may also be achieved by such substitutions, compared to the affinity of the parent peptide. In any event, such substitutions should employ amino acid residues or other molecular fragments chosen to avoid, for

10   example, steric and charge interference which might disrupt binding.

Amino acid substitutions are typically of single residues. Substitutions, deletions, insertions or any combination thereof may be combined to arrive at a final

15   peptide. Substitutional variants are those in which at least one residue of a peptide has been removed and a different residue inserted in its place. Such substitutions generally are made in accordance with the following Table 2 when it is desired to finely modulate the characteristics of the peptide.

TABLE 2

Original Residue	Exemplary Substitution
Ala	Ser
Arg	Lys, His
Asn	Gln
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Pro
His	Lys; Arg
Ile	Leu; Val
Leu	Ile; Val
Lys	Arg; His
Met	Leu; Ile
Phe	Tyr; Trp
Ser	Thr
Thr	Ser
Trp	Tyr; Phe
Tyr	Trp; Phe
Val	Ile; Leu

Substantial changes in function (e.g., affinity for MHC molecules or T cell receptors) are made by selecting substitutions that are less conservative than those in Table 2, i.e., selecting residues that differ more significantly in their effect on maintaining (a) the structure of the peptide backbone in the area of the substitution, for example as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site or (c) the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in peptide properties will be those in which (a) hydrophilic residue, e.g. seryl, is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a residue having an electropositive side chain, e.g., lysl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g. glutamyl or aspartyl; or (c) a residue having a bulky side chain, e.g. phenylalanine, is substituted for (or by) one not having a side chain, e.g., glycine.

The peptides may also comprise isosteres of two or more residues in the immunogenic peptide. An isostere as defined here is a sequence of two or more residues that can be substituted for a second sequence because the steric conformation of the first sequence fits a binding site specific for the second sequence. The term specifically includes peptide backbone modifications well known to those skilled in the art. Such modifications include modifications of the amide nitrogen, the α-carbon, amide carbonyl, complete replacement of the amide bond, extensions, deletions or backbone crosslinks. See, generally, Spatola, Chemistry and Biochemistry of Amino Acids, peptides and Proteins, Vol. VII (Weinstein ed., 1983).

Modifications of peptides with various amino acid mimetics or unnatural amino acids are particularly useful in increasing the stability of the peptide in vivo. Stability can be assayed in a number of ways. For instance, peptidases and various biological media, such as human plasma and serum, have been used to test stability. See, e.g., Verhoef et al.,

Eur. J. Drug Metab. Pharmacokin. 11:291-302 (1986). Half
life of the peptides of the present invention is conveniently
determined using a 25% human serum (v/v) assay. The protocol
is generally as follows. Pooled human serum (Type AB,

5   non-heat inactivated) is delipidated by centrifugation before
use. The serum is then diluted to 25% with RPMI tissue
culture media and used to test peptide stability. At
predetermined time intervals a small amount of reaction
solution is removed and added to either 6% aqueous

10   trichloracetic acid or ethanol. The cloudy reaction sample is
cooled (4°C) for 15 minutes and then spun to pellet the
precipitated serum proteins. The presence of the peptides is
then determined by reversed-phase HPLC using
stability-specific chromatography conditions.

15   The peptides of the present invention or analogs
thereof which have CTL stimulating activity may be modified to
provide desired attributes other than improved serum half
life. For instance, the ability of the peptides to induce CTL
activity can be enhanced by linkage to a sequence which

20   contains at least one epitope that is capable of inducing a T
helper cell response. Particularly preferred immunogenic
peptides/T helper conjugates are linked by a spacer molecule.
The spacer is typically comprised of relatively small, neutral
molecules, such as amino acids or amino acid mimetics, which

25   are substantially uncharged under physiological conditions.
The spacers are typically selected from, e.g., Ala, Gly, or
other neutral spacers of nonpolar amino acids or neutral polar
amino acids. It will be understood that the optionally
present spacer need not be comprised of the same residues and

30   thus may be a hetero- or homo-oligomer. When present, the
spacer will usually be at least one or two residues, more
usually three to six residues. Alternatively, the CTL peptide
may be linked to the T helper peptide without a spacer.

           The immunogenic peptide may be linked to the T helper
35   peptide either directly or via a spacer either at the amino or
carboxy terminus of the CTL peptide. The amino terminus of
either the immunogenic peptide or the T h lper peptide may be
acylated. Exemplary T helper peptides include tetanus toxoid

830-843, influenza 307-319, malaria circumsporozoite 382-398
and 378-389.

In some embodiments it may be desirable to include in
the pharmaceutical compositions of the invention at least one
5   component which primes CTL. Lipids have been identified as
agents capable of priming CTL _in vivo_ against viral antigens.
For example, palmitic acid residues can be attached to the
alpha and epsilon amino groups of a Lys residue and then
linked, e.g., via one or more linking residues such as Gly,
10   Gly-Gly-, Ser, Ser-Ser, or the like, to an immunogenic
peptide. The lipidated peptide can then be injected directly
in a micellar form, incorporated into a liposome or emulsified
in an adjuvant, e.g., incomplete Freund's adjuvant. In a
preferred embodiment a particularly effective immunogen
15   comprises palmitic acid attached to alpha and epsilon amino
groups of Lys, which is attached via linkage, e.g., Ser-Ser,
to the amino terminus of the immunogenic peptide.

As another example of lipid priming of CTL responses,
E. coli lipoproteins, such as
20   tripalmitoyl-S-glycerylcysteinlyseryl-serine (P$_3$CSS) can be
used to prime virus specific CTL when covalently attached to
an appropriate peptide. See, Deres et al., _Nature_ 342:561-564
(1989), incorporated herein by reference. Peptides of the
invention can be coupled to P$_3$CSS, for example, and the
25   lipopeptide administered to an individual to specifically
prime a CTL response to the target antigen. Further, as the
induction of neutralizing antibodies can also be primed with
P$_3$CSS conjugated to a peptide which displays an appropriate
epitope, the two compositions can be combined to more
30   effectively elicit both humoral and cell-mediated responses to
infection.

In addition, additional amino acids can be added to
the termini of a peptide to provide for ease of linking
peptides one to another, for coupling to a carrier support, or
35   larger peptide, for modifying the physical or chemical
properties of the peptide or oligopeptide, or the like. Amino
acids such as tyrosine, cysteine, lysine, glutamic or aspartic
acid, or the like, can be introduced at the C- or N-terminus

of the peptide or oligopeptide. Modification at the C terminus in some cases may alter binding characteristics of the peptide. In addition, the peptide or oligopeptide sequences can differ from the natural sequence by being

5   modified by terminal-NH_2 acylation, e.g., by alkanoyl (C_1-C_{20}) or thioglycolyl acetylation, terminal-carboxyl amidation, e.g., ammonia, methylamine, etc. In some instances these modifications may provide sites for linking to a support or other molecule.

10   The peptides of the invention can be prepared in a wide variety of ways. Because of their relatively short size, the peptides can be synthesized in solution or on a solid support in accordance with conventional techniques. Various automatic synthesizers are commercially available and can be

15   used in accordance with known protocols. See, for example, Stewart and Young, <u>Solid Phase Peptide Synthesis</u>, 2d. ed., Pierce Chemical Co. (1984), <u>supra</u>.

      Alternatively, recombinant DNA technology may be employed wherein a nucleotide sequence which encodes an

20   immunogenic peptide of interest is inserted into an expression vector, transformed or transfected into an appropriate host cell and cultivated under conditions suitable for expression. These procedures are generally known in the art, as described generally in Sambrook et al., <u>Molecular Cloning, A Laboratory</u>

25   <u>Manual</u>, Cold Spring Harbor Press, Cold Spring Harbor, New York (1982), which is incorporated herein by reference. Thus, fusion proteins which comprise one or more peptide sequences of the invention can be used to present the appropriate T cell epitope.

30   As the coding sequence for peptides of the length contemplated herein can be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci et al., <u>J. Am. Chem. Soc.</u> 103:3185 (1981), modification can be made simply by substituting the appropriate base(s) for those

35   encoding the native peptide sequence. The coding sequence can then be provided with appropriate linkers and ligated into expression vectors commonly available in the art, and the vectors used to transform suitable hosts to produce the

desired fusion protein. A number of such vectors and suitable host systems are now available. For expression of the fusion proteins, the coding sequence will be provided with operably linked start and stop codons, promoter and terminator regions and usually a replication system to provide an expression vector for expression in the desired cellular host. For example, promoter sequences compatible with bacterial hosts are provided in plasmids containing convenient restriction sites for insertion of the desired coding sequence. The resulting expression vectors are transformed into suitable bacterial hosts. Of course, yeast or mammalian cell hosts may also be used, employing suitable vectors and control sequences.

The peptides of the present invention and pharmaceutical and vaccine compositions thereof are useful for administration to mammals, particularly humans, to treat and/or prevent viral infection and cancer. Examples of diseases which can be treated using the immunogenic peptides of the invention include prostate cancer, hepatitis B, hepatitis C, AIDS, renal carcinoma, cervical carcinoma, lymphoma, CMV and condlyloma acuminatum.

For pharmaceutical compositions, the immunogenic peptides of the invention are administered to an individual already suffering from cancer or infected with the virus of interest. Those in the incubation phase or the acute phase of infection can be treated with the immunogenic peptides separately or in conjunction with other treatments, as appropriate. In therapeutic applications, compositions are administered to a patient in an amount sufficient to elicit an effective CTL response to the virus or tumor antigen and to cure or at least partially arrest symptoms and/or complications. An amount adequate to accomplish this is defined as "therapeutically effective dose." Amounts effective for this use will depend on, e.g., the peptide composition, the manner of administration, the stage and severity of the disease being treated, the weight and gen ral state of health of the patient, and the judgment of the prescribing physician, but generally range for the initial

immunization (that is for therapeutic or prophylactic administration) from about 1.0 μg to about 5000 μg of peptide for a 70 kg patient, followed by boosting dosages of from about 1.0 μg to about 1000 μg of peptide pursuant to a

5   boosting regimen over weeks to months depending upon the patient's response and condition by measuring specific CTL activity in the patient's blood. It must be kept in mind that the peptides and compositions of the present invention may generally be employed in serious disease states, that is,

10   life-threatening or potentially life threatening situations. In such cases, in view of the minimization of extraneous substances and the relative nontoxic nature of the peptides, it is possible and may be felt desirable by the treating physician to administer substantial excesses of these peptide

15   compositions.

For therapeutic use, administration should begin at the first sign of viral infection or the detection or surgical removal of tumors or shortly after diagnosis in the case of acute infection. This is followed by boosting doses until at

20   least symptoms are substantially abated and for a period thereafter. In chronic infection, loading doses followed by boosting doses may be required.

Treatment of an infected individual with the compositions of the invention may hasten resolution of the

25   infection in acutely infected individuals. For those individuals susceptible (or predisposed) to developing chronic infection the compositions are particularly useful in methods for preventing the evolution from acute to chronic infection. Where the susceptible individuals are identified prior to or

30   during infection, for instance, as described herein, the composition can be targeted to them, minimizing need for administration to a larger population.

The peptide compositions can also be used for the treatment of chronic infection and to stimulate the immune

35   system to eliminate virus-infected cells in carriers. It is important to provide an amount of immuno-potentiating peptide in a formulation and mode of administration sufficient to effectively stimulate a cytotoxic T cell response. Thus, for

treatment of chronic infection, a representative dose is in the range of about 1.0 μg to about 5000 μg, preferably about 5 μg to 1000 μg for a 70 kg patient per dose. Immunizing dos s followed by boosting doses at established intervals, e.g.,

5   from one to four weeks, may be required, possibly for a prolonged period of time to effectively immunize an individual. In the case of chronic infection, administration should continue until at least clinical symptoms or laboratory tests indicate that the viral infection has been eliminated or

10   substantially abated and for a period thereafter.

The pharmaceutical compositions for therapeutic treatment are intended for parenteral, topical, oral or local administration. Preferably, the pharmaceutical compositions are administered parenterally, e.g., intravenously,

15   subcutaneously, intradermally, or intramuscularly. Thus, the invention provides compositions for parenteral administration which comprise a solution of the immunogenic peptides dissolved or suspended in an acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers may be used,

20   e.g., water, buffered water, 0.8% saline, 0.3% glycine, hyaluronic acid and the like. These compositions may be sterilized by conventional, well known sterilization techniques, or may be sterile filtered. The resulting aqueous solutions may be packaged for use as is, or lyophilized, the

25   lyophilized preparation being combined with a sterile solution prior to administration. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions, such as pH adjusting and buffering agents, tonicity adjusting agents, wetting

30   agents and the like, for example, sodium acetate, sodium lactate, sodium chloride, potassium chloride, calcium chloride, sorbitan monolaurate, triethanolamine oleate, etc.

The concentration of CTL stimulatory peptides of the invention in the pharmaceutical formulations can vary widely,

35   i.e., from less than about 0.1%, usually at or at least about 2% to as much as 20% to 50% or more by weight, and will be selected primarily by fluid volumes, viscosities, etc., in

accordance with the particular mode of administration selected.

The peptides of the invention may also be administered via liposomes, which serve to target the peptides to a particular tissue, such as lymphoid tissue, or targeted selectively to infected cells, as well as increase the half-life of the peptide composition. Liposomes include emulsions, foams, micelles, insoluble monolayers, liquid crystals, phospholipid dispersions, lamellar layers and the like. In these preparations the peptide to be delivered is incorporated as part of a liposome, alone or in conjunction with a molecule which binds to, e.g., a receptor prevalent among lymphoid cells, such as monoclonal antibodies which bind to the CD45 antigen, or with other therapeutic or immunogenic compositions. Thus, liposomes either filled or decorated with a desired peptide of the invention can be directed to the site of lymphoid cells, where the liposomes then deliver the selected therapeutic/immunogenic peptide compositions. Liposomes for use in the invention are formed from standard vesicle-forming lipids, which generally include neutral and negatively charged phospholipids and a sterol, such as cholesterol. The selection of lipids is generally guided by consideration of, e.g., liposome size, acid lability and stability of the liposomes in the blood stream. A variety of methods are available for preparing liposomes, as described in, e.g., Szoka et al., _Ann. Rev. Biophys. Bioeng_. 9:467 (1980), U.S. Patent Nos. 4,235,871, 4,501,728, 4,837,028, and 5,019,369, incorporated herein by reference.

For targeting to the immune cells, a ligand to be incorporated into the liposome can include, e.g., antibodies or fragments thereof specific for cell surface determinants of the desired immune system cells. A liposome suspension containing a peptide may be administered intravenously, locally, topically, etc. in a dose which varies according to, _inter_ _alia_, the manner of administration, the peptide being delivered, and the stage of the disease being treated.

For solid compositions, conventional nontoxic solid carriers may be used which include, for example,

pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharin, talcum, cellulose, glucose, sucrose, magnesium carbonate, and the like. For oral administration, a pharmaceutically acceptable nontoxic

5   composition is formed by incorporating any of the normally employed excipients, such as those carriers previously listed, and generally 10-95% of active ingredient, that is, one or more peptides of the invention, and more preferably at a concentration of 25%-75%.

10   For aerosol administration, the immunogenic peptides are preferably supplied in finely divided form along with a surfactant and propellant. Typical percentages of peptides are 0.01%-20% by weight, preferably 1%-10%. The surfactant must, of course, be nontoxic, and preferably soluble in the

15   propellant. Representative of such agents are the esters or partial esters of fatty acids containing from 6 to 22 carbon atoms, such as caproic, octanoic, lauric, palmitic, stearic, linoleic, linolenic, olesteric and oleic acids with an aliphatic polyhydric alcohol or its cyclic anhydride. Mixed

20   esters, such as mixed or natural glycerides may be employed. The surfactant may constitute 0.1%-20% by weight of the composition, preferably 0.25-5%. The balance of the composition is ordinarily propellant. A carrier can also be included, as desired, as with, e.g., lecithin for intranasal

25   delivery.

        In another aspect the present invention is directed to vaccines which contain as an active ingredient an immunogenically effective amount of an immunogenic peptide as described herein. The peptide(s) may be introduced into a

30   host, including humans, linked to its own carrier or as a homopolymer or heteropolymer of active peptide units. Such a polymer has the advantage of increased immunological reaction and, where different peptides are used to make up the polymer, the additional ability to induce antibodies and/or CTLs that

35   react with different antigenic determinants of the virus or tumor cells. Useful carriers are well known in the art, and include, e.g., thyroglobulin, albumins such as human serum albumin, tetanus toxoid, polyamino acids such as

poly(lysine:glutamic acid), influenza, hepatitis B virus core
protein, hepatitis B virus recombinant vaccine and the like.
The vaccines can also contain a physiologically tolerable
(acceptable) diluent such as water, phosphate buffered saline,
or saline, and further typically include an adjuvant.
Adjuvants such as incomplete Freund's adjuvant, aluminum
phosphate, aluminum hydroxide, or alum are materials well
known in the art. And, as mentioned above, CTL responses can
be primed by conjugating peptides of the invention to lipids,
such as P_3CSS. Upon immunization with a peptide composition
as described herein, via injection, aerosol, oral, transdermal
or other route, the immune system of the host responds to the
vaccine by producing large amounts of CTLs specific for the
desired antigen, and the host becomes at least partially
immune to later infection, or resistant to developing chronic
infection.

Vaccine compositions containing the peptides of the
invention are administered to a patient susceptible to or
otherwise at risk of viral infection or cancer to elicit an
immune response against the antigen and thus enhance the
patient's own immune response capabilities. Such an amount is
defined to be an "immunogenically effective dose." In this
use, the precise amounts again depend on the patient's state
of health and weight, the mode of administration, the nature
of the formulation, etc., but generally range from about 1.0
μg to about 5000 μg per 70 kilogram patient, more commonly
from about 10 μg to about 500 μg mg per 70 kg of body weight.

In some instances it may be desirable to combine the
peptide vaccines of the invention with vaccines which induce
neutralizing antibody responses to the virus of interest,
particularly to viral envelope antigens.

For therapeutic or immunization purposes, the peptides
of the invention can also be expressed by attenuated viral
hosts, such as vaccinia or fowlpox. This approach involves
the use of vaccinia virus as a vector to express nucleotide
sequences that encode the peptides of the invention. Upon
introduction into an acutely or chronically infected host or
into a non-infected host, the recombinant vaccinia virus

expresses the immunogenic peptide, and thereby elicits a host CTL response. Vaccinia vectors and methods useful in immunization protocols are described in, e.g., U.S. Patent No. 4,722,848, incorporated herein by reference. Another vector

5 is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover et al. (<u>Nature</u> 351:456-460 (1991)) which is incorporated herein by reference. A wide variety of other vectors useful for therapeutic administration or immunization of the peptides of the invention, e.g., <u>Salmonella</u> <u>typhi</u>

10 vectors and the like, will be apparent to those skilled in the art from the description herein.

Antigenic peptides may be used to elicit CTL <u>ex vivo</u>, as well. The resulting CTL, can be used to treat chronic infections (viral or bacterial) or tumors in patients that do

15 not respond to other conventional forms of therapy, or will not respond to a peptide vaccine approach of therapy. <u>Ex vivo</u> CTL responses to a particular pathogen (infectious agent or tumor antigen) are induced by incubating in tissue culture the patient's CTL precursor cells (CTLp) together with a source of

20 antigen-presenting cells (APC) and the appropriate immunogenic peptide. After an appropriate incubation time (typically 1-4 weeks), in which the CTLp are activated and mature and expand into effector CTL, the cells are infused back into the patient, where they will destroy their specific target cell

25 (an infected cell or a tumor cell).

The peptides may also find use as diagnostic reagents. For example, a peptide of the invention may be used to determine the susceptibility of a particular individual to a treatment regimen which employs the peptide or related

30 peptides, and thus may be helpful in modifying an existing treatment protocol or in determining a prognosis for an affected individual. In addition, the peptides may also be used to predict which individuals will be at substantial risk for developing chronic infection.

35 The following example is offered by way of illustration, not by way of limitation.

Example 1

Class I antigen isolation was carried out as described in the parent applications. Naturally processed peptides were then Isolated and sequenced as described there. An allele-

5 specific motif and algorithms were determined and quantitative binding assays were carried.

Using the motifs identified above for HLA-A2.1 allele amino acid sequences from a tumor-related proteins, Melanoma Antigen-1,-2, and -3 (MAGE-1, -2, and -3), were analyzed for

10 the presence of these motifs. Sequences for the target antigen are obtained from the GenBank data base (Release No. 71.0; 3/92). The identification of motifs is done using the "FINDPATTERNS" program (Devereux et al., Nucleic Acids Research 12:387-395 (1984)).

15 Other viral and tumor-related proteins can also be analyzed for the presence of these motifs. The amino acid sequence or the nucleotide sequence encoding products is obtained from the GenBank database in the cases of Human Papilloma Virus (HPV), Prostate Specific antigen (PSA), p53

20 oncogene, Epstein Barr Nuclear Antigen-1 (EBNA-1), and c-erb2 oncogene (also called HER-2/neu).

In the cases of Hepatitis B Virus (HBV), Hepatitis C Virus (HCV), and Human Immunodeficiency Virus (HIV) several strains/isolates exist and many sequences have been placed in

25 GenBank.

For HBV, binding motifs are identified for the adr, adw and ayw types. In order to avoid replication of identical sequences, all of the adr motifs and only those motifs from adw and ayw that are not present in adr are added to the list

30 of peptides.

In the case of HCV, a consensus sequence from residue 1 to residue 782 is derived from 9 viral isolates. Motifs are identified on those regions that have no or very little (one residue) variation between the 9 isolates. The sequences of

35 residues 783 to 3010 from 5 viral isolates were also analyzed. Motifs common to all the isolates are identified and added to the peptide list.

Finally, a consensus sequence for HIV type 1 for North American viral isolates (10-12 viruses) was obtained from the Los Alamos National Laboratory database (May 1991 release) and analyzed in order to identify motifs that are constant

5   throughout most viral isolates. Motifs that bear a small degree of variation (one residue, in 2 forms) were also added to the peptide list.

Tables 3 and 4 provide the results of these searches. Binding affinities are expressed as percentage of binding

10   compared to standard peptide in the assays as described in the parent applications are presented.

The above examples are provided to illustrate the invention but not to limit its scope. Other variants of the invention will be readily apparent to one of ordinary skill in

15   the art and are encompassed by the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference.

Table 3

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
ALEAQQEAL	9	1		15	2.1		<0.0003			
ILESLFRAV	9	1		93	2.1		0.0004			
VITKKVADL	9	1		101	2.1		<0.0003			
CLGLSYDGL	9	1/3		174	2.1		0.0004			
QIMPKTGFL	9	1		187	2.1		0.0007			
SLHCKPEEAL	10	1		7	2.1		0.0002			
PLVLGTLEEV	10	1		37	2.1		0.0008			
CILESLFRAV	10	1		92	2.1		0.0003			
AVITKKVADL	10	1		100	2.1		0			
VITKKVADLV	10	1		101	2.1		0			
LLKYRAREPV	10	1/3		114	2.1		0			
EIFGKASESL	10	1		142	2.1		0			
CLGLSYDGLL	10	1/3		174	2.1		0			
AISRKMVEL	9	2		101	2.1		0.0003			
KMVELVHFL	9	2		105	2.1		0.16			
MVELVHFLL	9	2		106	2.1		0.0031			
DLQQSLRVL	9	2		143	2.1		0			
SLRVLAAGL	9	2		147	2.1		0.0001			
ALSRKVAEL	9	3		101	2.1		0.0050			
HLYIFATCL	9	3		167	2.1		0.0003			
YIFATCLGL	9	3		169	2.1		0.018			
QIMPKAGLL	9	3		187	2.1		0			

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
AISRKMVELV	10	2		101	2.1		0			
MVELVHFLLL	10	2		106	2.1		0.0017			
KLPGLLSRDL	10	2		135	2.1		0			
LLSRDLQQSL	10	2		139	2.1		0.0007			
SLPTTMNYPL	10	3		63	2.1		0.0035			
DLESEFQAAL	10	3		93	2.1		0.0001			
ALSRKVAELV	10	3		101	2.1		0.0001			
KVAELVHFLL	10	3		105	2.1		0.012			
VIFSKASSSL	10	3		142	2.1		0			
SLQLVFGIEL	10	3		150	2.1		0.0049			
LMEVDPIGHL	10	3		159	2.1		0.0005			
FLIIVLVMI	9	1		194	2.1		0.0005			
GLLGDNQIM	9	1		181	2.1		0.0051			
SLHCKPEEA	9	1		7	2.1		0.013	<0.0002	0	
ALGLVCVQA	9	1		22	2.1		0.015	<0.0002	<0.0002	
CKPEEALEA	9	1		10	Random		<0.0002			
QQEALGLVC	9	1		19	Random		<0.0002			
VQAATSSSS	9	1		28	Random		<0.0002			
PLVLGTLEE	9	1		37	Random		<0.0002			
VPTAGSTDP	9	1		46	Random		<0.0002			
PQSPQGASA	9	1		55	Random		<0.0002			
FPTTINFTR	9	1		64	Random		<0.0002			

Table 3
Page 2 of 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
QRQPSEGSS	9	1		73	Random		<0.0002			
SREEEGPST	9	1		82	Random		<0.0002			
AVITKKVAD	9	1		100	Random		<0.0002			
EMLESVIKN	9	1		127	Random		<0.0002			0
YKHCFPEIF	9	1		136	Random		<0.0002			
GKASESLQL	9	1		145	Random		<0.0002			
VFGIDVKEA	9	1		154	Random		<0.0002	<0.0002	0	
DPTGHSYVL	9	1		163	Random		<0.0002			
VTCLGLSYD	9	1		172	Random		<0.0002			
PKTGFLIIV	9	1		190	Random		<0.0002			
LVMIAMEGG	9	1		199	Random		<0.0002			
HAPEEEIWE	9	1		208	Random		<0.0002			
ELSVMEVYD	9	1		217	Random		<0.0002			
GREHSAYGE	9	1		226	Random		<0.0002			
PRKLLTQDL	9	1		235	Random		0.0002			
VQEKYLEYG	9	1		244	Random		<0.0002			
RCRTVIPHA	9	1		253	Random		<0.0002			
MSSCGVQGP	9	1		262	Random		<0.0002			
ILESLFRAVI	10	1		93	2.1		0.0002			
FLIIVLVMIA	10	1		194	2.1		0.0003	0.0093	0.0030	
LVFGIDVKEA	10	1		153	2.1		0.0002	<0.0002	0	
EVYDGREHSA	10	1		222	2.1		0	<0.0002	0	

Table 3
Page 3 of 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
GVQGPSLKPA	10	1		266	2.1		0.0001			
QLVFGIDV	8	1		152	2.1		0			
KLLTQDLV	8	1		237	2.1		0.0004			
GLLGDNQI	8	1		181	2.1		0			
DLVGFLLL	8	1		108	2.1		0			
GLSYDGLL	8	1		176	2.1		0.0001			
DLVQEKYL	8	1		242	2.1		0			
LLGDNQIM	8	1		182	2.1		0			
FLIIVLVM	8	1		194	2.1		0			
ALEAQQEA	8	1		15	2.1		0			
TLEEVPTA	8	1		42	2.1		0			
IMPKTGFL	8	1		188	2.1		0.0001			
PVTKAEML	8	1		122	2.1		0			
IVLVMIAM	8	1		197	2.1		0.0001			
AVITKKVA	8	1		100	2.1		0			
EIWEELSV	8	1		213	2.1		0			
LIIVLVMI	8	1		195	2.1		0.0001			
IIVLVMIA	8	1		196	2.1		0.0002			
SLFRAVITKKV	11	1		96	2.1		0.0001			
LLLKYRAREPV	11	1		113	2.1		0.0001			
YLEYGRCRTVI	11	1		248	2.1		0.0006			
ALEAQQEALGL	11	1		15	2.1		0.0001			

Table 3
Page 4 of 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
FLIIVLVMIAM	11	1		194	2.1		0.0041			
VLGTLEEVPTA	11	1		39	2.1		0.0002			
QLVFGIDVKEA	11	1		152	2.1		0.0001			
AVITKKVADLV	11	1		100	2.1		0			
PVTKAEMLESV	11	1		122	2.1		0			
KVADLVGFLLL	11	1		105	2.1		0.020			
GVQGPSLKPAM	11	1		266	2.1		0			
LVGFLLLKYRA	11	1		109	2.1		0.0004			
LVMIAMEGGHA	11	1		199	2.1		0.0005			
CILESLFRAVI	11	1		92	2.1		0.0030			
EALEAQQEA	9	1		14	2.1		0	<0.0002	0	
EAQQEALGL	9	1		17	2.1		0			<0.0002
AATSSSSPL	9	1		30	2.1		0			<0.0002
ATSSSSPLV	9	1		31	2.1		0.0007			
GTLEEVPTA	9	1		41	2.1		0.013	<0.0002	0	
GASAFPTTI	9	1		60	2.1		0			<0.0002
STSCILESL	9	1		89	2.1		0.0002			
RAVITKKVA	9	1		99	2.1		0	<0.0002	0	
ITKKVADLV	9	1		102	2.1		0			
RAREPVTKA	9	1		118	2.1		0			
KAEMLESVI	9	1		125	2.1		0			<0.0002
KASESLQLV	9	1		146	2.1		0.0009			

Table 3
Page 5 of 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
PTGHSYLV	9	1		164	2.1		0			
KTGFLIIVL	9	1		191	2.1		0.0006			
LIIVLVMIA	9	1		195	2.1		0	0.0022	0.0006	
IIVLVMIAM	9	1		196	2.1		0.0007			
MIAMEGGHA	9	1		201	2.1		0.0005	<0.0002	0.0002	
EIWEELSVM	9	1		213	2.1		0			
SAYGEPRKL	9	1		230	2.1		0.0002			<0.0002
YLEYGRCRT	9	1		248	2.1		0			
EALGLVCVQA	10	1		21	2.1		0.0005	<0.0002	0	
QAATSSSSPL	10	1		29	2.1		0			<0.0002
VTKAEMLESV	10	1		123	2.1		0			
EADPTGHSYV	10	1		161	2.1		0			
VLGTLEEVPT	10	1		39	2.1		0.0004			
SAFPTTINFT	10	1		62	2.1		0			
GIDVKEADPT	10	1		156	2.1		0			
PTGHSYLVT	10	1		164	2.1		0			
FLWGPRALA	9	1	new	265	2.1		0.042	0.0017	0	
LAETSYVKV	9	1	new	272	2.1		0			
YVKVLEYVI	9	1	new	277	2.1		0.0002			
RVRFFFPSL	9	1	new	290	2.1		0.0001			
LAETSYVKVL	10	1	new	272	2.1		0			<0.0002
VLEYVIKVSA	10	1	new	280	2.1		0.0002	0.0002	0	

Table 3
Page 6 of 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
AALREEEGV	10	1	new	301	2.1		0			
SMHCKPEEV	9	1	new (a)	7	2.1		0.018			
AMGLVCVQV	9	1	new (a)	22	2.1		0.012			
LMLGTLEEV	9	1	new (a)	38	2.1		0.13			
LQLVFGIDV	9	1	new	151	2.1		0.0004			
GLSYDGLLG	9	1	new	176	2.1		0			
GLSYDGLLV	9	1	new (a)	176	2.1		0.0047			
LLGDNQIMP	9	1	new	182	2.1		0.0001			
LLGDNQIMV	9	1	new (a)	182	2.1		0.043			
WEELSVMEV	9	1	new	215	2.1		0			
WMELSVMEV	9	1	new (a)	215	2.1		0.041			
RKLLTQDLV	9	1	new	236	2.1		0			
YEFLWGPRA	9	1	new	262	2.1		0			
YMFLWGPRV	9	1	new (a)	262	2.1		0.22			
AATSSSSPLV	10	1	new	30	2.1		0			
ATSSSSPLVL	10	1	new	31	2.1		0			
KMADLVGFLV	10	1	new (a)	105	2.1		1.5			
VADLVGFLLL	10	1	new	106	2.1		0.0008			0.0003
SESLQLVFGI	10	1	new	148	2.1		0			
VMVTCLGLSV	10	1	new (a)	170	2.1		0.30			
QIMPKTGFLI	10	1	new	187	2.1		0.0009			
QMPKTGFLIV	10	1	new (a)	187	2.1		0.050			

Table 3
Page 7 of 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
KTGFLIIVLV	10	1	new	191	2.1		0.0012			
LIIVLVMIAM	10	1	new	195	2.1		0.0003			
VMIAMEGGHV	10	1	new (a)	200	2.1		0.053			
SAYGEPRKLL	10	1	new	230	2.1		0			0.0008
ALAETSYVKVL	11	1 N		270	2.1		0.012			
KMVELVHFLLL	11	2		52	2.1		0.67			
ELMEVDPIGHL	11	3		105	2.1		0.026			
HLYIFATCLGL	11	3		114	2.1		0.041			
LLLKYRAREPV	11	3		60	2.1		0.0001			
QLVFGIELMEV	11	3		99	2.1		0.34			
IMPKAGLLIIV	11	3		135	2.1		0.013			
VLVTCLGLSYDGL	13	1 n	E6	170	2.1		0.0017			
KLLTQDLVQEKYL	13	1 n	E6	237	2.1		0.0060			
DLVQEKYLEYRQV	13	1 n	E6	242	2.1		0			
SLFRAVITKKVADLV	15	1 n	POL	96	2.1		0.0004			
DLESEFQAAISRKMV	15	2	POL	40	2.1		0			
MLGSVVGNWQYFFPV	15	3	POL	75	2.1		0.012			
GASSFSTTI	9	2		60	2.1		0			0.0002
DLESEFQAA	9	2,3		93	2.1		0			
QAAISRKMV	9	2		99	2.1		0			
KAEMLESVL	9	2		125	2.1		0			0
KASEYLQLV	9	2		146	2.1		0.011			

Table 3
Page 8 of 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
QLVFGIEVV	9	2		152	2.1		0.0038			
VVPISHLYI	9	2		162	2.1		0.0002			
PISHLYILV	9	2		164	2.1		0.0005			
HLYILVTCL	9	2		167	2.1		0.0034			
YILVTCLGL	9	2		169	2.1		0.0014			
GLLGDNQVM	9	2		181	2.1		0.0038			
QVMPKTGLL	9	2		187	2.1		0			
VMPKTGLLI	9	2		188	2.1		0.0010			0.230
KTGLLIIVL	9	2		191	2.1		0.0002			
GLLIIVLAI	9	2,3		193	2.1		0.0002			
LLIIVLAII	9	2,3		194	2.1		0.0001			
LIIVLAIIA	9	2,3		195	2.1		0.0008			
IIVLAIIAI	9	2		196	2.1		0.0009			
IIAIEGDCA	9	2		201	2.1		0			
GASSLPTTM	9	3		60	2.1		0			0.0010
QAALSRKVA	9	3		99	2.1		0			
VAELVHFLL	9	3		106	2.1		0			0.039
KAEMLGSVV	9	3		125	2.1		0			
KASSSLQLV	9	3		146	2.1		0.0005			
QLVFGIELM	9	3		152	2.1		0.0010			
PIGHLYIFA	9	3		164	2.1		0			
IMPKAGLLI	9	3		188	2.1		0.0064			

Table 3
Page 9 of 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
KAGLLIIVL	9	3		191	2.1		0.0002			0
IIAREGDCA	9	3		201	2.1		0			
EALEAQQEAL	10	1	new	14	2.1		0			0
EAQQEALGLV	10	1	new	17	2.1		0			
DLESEFQAAI	10	2		93	2.1		0			
AAISRKMVEL	10	2		100	2.1		0			0
VIFSKASEYL	10	2		142	2.1		0.0014			
YLQLVFGIEV	10	2		150	2.1		0.37			
LVFGIEVVEV	10	2		153	2.1		0.012			
GIEVVEVVPI	10	2		156	2.1		<0.0002			
VVEVVPISHL	10	2		159	2.1		<0.0002			
EVVPISHLYI	10	2		161	2.1		<0.0002			
VVPISHLYIL	10	2		162	2.1		0.0002			
PISHLYILVT	10	2		164	2.1		0.0003			
QVMPKTGLLI	10	2		187	2.1		0.0002			
VMPKTGLLII	10	2		188	2.1		0.0009			0.058
KTGLLIIVLA	10	2		191	2.1		<0.0002			
GLLIIVLAII	10	2,3		193	2.1		0.0005			
LLIIVLAIIA	10	2,3		194	2.1		<0.0002			
LIIVLAIIAI	10	2		195	2.1		0.0013			
AIIAIEGDCA	10	2		200	2.1		0.0023			
AALSRKVAEL	10	3		100	2.1		0.0007			0

Table 3

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
VAELVHFLLL	10	3		106	2.1		0.0009			0.0018
VTKAEMLGSV	10	3		123	2.1		<0.0002			
GIELMEVDPI	10	3		156	2.1		<0.0002			
EVDPIGHLYI	10	3		161	2.1		<0.0002			
PIGHLYIFAT	10	3		164	2.1		0.0003			
QIMPKAGLLI	10	3		187	2.1		0.0006			
IMPKAGLLII	10	3		188	2.1		0.0015			
KAGLLIIVLA	10	3		191	2.1		<0.0002			
AIIAREGDCA	10	3		200	2.1		<0.0002			
FLWGPRALI	9	2		271	A02					
GLEARGEAL	9	3		15	A02					
EARGEALGL	9	3		17	A02					
ALGLVGAQA	9	3		22	A02/A03					
GLVGAQAPA	9	3		24	A02/A03					
LVGAQAPAT	9	3		25	A02					
PATEEQEAA	9	3		31	A02/A03					
EAASSSSTL	9	3		37	A02					
AASSSSTLV	9	3		38	A02					
LVEVTLGEV	9	3		45	A02					
EVTLGEVPA	9	3		47	A02/A03					
VTLGEVPAA	9	3		48	A02/A03					
KIWEELSVL	9	3		220	A02					

Table 3
Page 11 of 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
SILGDPKKL	9	3		237	A02					
ILGDPKKLL	9	3		238	A02					
FLWGPRALV	9	3		271	A02					
RALVETSYV	9	3		276	A02					
LVETSYVKV	9	3		278	A02					
YVKVLHHMV	9	3		283	A02					
KVLHHMVKI	9	3		285	A02					
EARGEALGLV	10	3		17	A02					
EALGLVGAQA	10	3		21	A02/A03					
GLVGAQAPAT	10	3		24	A02					
QAPATEEQEA	10	3		29	A02/A03					
EAASSSSTLV	10	3		37	A02					
TLVEVTLGEV	10	3		44	A02					
EVTLGEVPAA	10	3		47	A02/A03					
EVFEGREDSI	10	3		229	A02					
SILGDPKKLL	10	3		237	A02					
ILGDPKKLLT	10	3		238	A02					
ALVETSYVKV	10	3		277	A02					
LVETSYVKVL	10	3		278	A02					
MVKISGGPHI	10	3		290	A02					
LVLGTLEEV	9	1		38	2.1	<0.0006	0.032	0	0	0.0003
KVADLVGFLL	10	1		105		0.0005	0.041	0.0039	0.0030	0.0070

Table 3
Page 12 of 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
LVFGIELMEV	10	3		153	2.1		0.17			
ILLWQPIPV	9	3				<0.0007	1.4	0.0048	0.0048	0
EVDPIGHLY	9	3				3.7			0.0022	0.0043
KMVELVHFL	9	2				<0.0007	0.13	0.0007	0	0.0008
KMVELVHFLL	10	2		105		<0.0008	0.071	0.0004	0.0001	0
LVFGIELMEV	10	3				0.0030	0.065	0.0007	0	0.0005
KVAELVHFL	9	3		105	2.1	0	0.073	0.011	0.0047	0
CILESLFRA	9	1		92	2.1	0.0001	0.073	0	0.0002	0
VMIAMEGGHA	10	1		200	2.1	<0.00008	0.0023	0	0	0
MLESVIKNYK	10	1				0	0	0.034	0.0045	0
ETSYVKVLEY	10	1				0.075	0	0.0009	0.0004	0
KVLEYVIKV	9	1	new	279	2.1	<0.0005	0.095	0.022	0.015	0
FLWGPRALA	9	1				<0.0006	0.027	0.0015	0	0
ALREEEEGV	9	1		302	2.1	<0.0006	0.0056	0	0	0
ALAETSYVKV	10	1		271	2.1	<0.0007	0.017	0.0011	0.0029	0
YVIKVSARV	9	1		283	2.1	0.0005	0.018	0	0	0
RALAETSYV	9	1		270	2.1	<0.0006	0.014	0.0003	0.0005	0
ALAETSYVK	9	1				<0.0006	0.0002	0.17	0.39	0
VLGTLEEV	8	1		39	2.1	<0.0007	0.0088	0	0	0
SLQLVFGI	8	1		150	2.1	<0.0007	0.0094	0	0.0001	0
ILESLFRA	8	1		93	2.1	<0.0004	0.0017	0.0003	0	0.0001
FLLLKYRA	8	1		112	2.1	0.0036	0.0007	0.0003	0.0001	0

Table 3
Page 13 of 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
GLVCVQAA	8	1		24	2.1	0.0016	0.0008	0.0008	0	0
VLVTCLGL	8	1		170	2.1	<0.0007	0.0010	0.0001	0	0
KVADLVGFL	9	1		105	2.1	<0.0008	0.0091	0.0013	0.0005	0
YVLVTCLGL	9	1		169	2.1					
IMPKTGFLI	9	1		188	2.1	<0.0008	0.0035	0	0	3.2
GLLGDNQIM	9	1			A2.1	<0.0008	0.0054	0	0	0.0002
GLVCVQAAT	9	1		24	2.1	0.0030	0.0007	0.0026	0	0.0001
VADLVGFLI	9	1		106	2.1	0.032	0.0011	0.0054	0.0008	0.0007
YLEYGRCRTV	10	1		248	2.1	0.0008	0.0097	0.0001	0	0
SLQIVFGIDV	10	1		150	2.1	0.0028	0.0047	0.0013	0.0001	0.0001
IMPKTGFLII	10	1		188	2.1	<0.0008	0.0007	0	0	0.050
ALGLVCVQAA	10	1		22	A2.1	0.0011	0.0002	0.0003	0	0
EIWEELSVMEV	11	1		213	A2.1	0.0007	0.013	0.0001	0.0001	0
FLIIVLVMIAM	11	1			A2.1	0.023	0.0031	0.016	0.0014	0.0011
VIPHAMSSCGV	11	1		257	2.1	<0.0009	1.4	0	0	0
CILESCFRAVI	11	1			A2.1	0.079	0.0017	0.058	0.0005	0.0008
QIMPKTGFLII	11	1		187	2.1	<0.0009	0.0003	0	0	0.0030
GFLLLKYRA	9	1						0.0004	0.0002	
CFPEIFGKA	9	1						0	0	
FFFPSLREA	9	1						0	0	
FFPSLREAA	9	1						0	0	
RSLHCKPEEA	10	1						0.0001	0.0008	

Table 3

Page 14 of 15

Sequence	AA	Mage Strain	Mol.	Pos.	Motif	A1	A2.1	A3.2	A11	A24
EFLWGPRALA	10	1						0	0	
RFFFPSLREA	10	1						0.0004	0	
FFFPSLREAA	10	1						0	0	

F:\PERFSOLDOCS\03939202.1

Table 3
Page 15 of 15

Table 4

Sequence	Antigen	Strain	Molecule	Position	Motif	A1 Binding	A2 Binding	A3 Binding	A11 Binding	A24 Binding	Max. Binding
ALFLGFLGAA	HIV	MN	gp160	518	A02		0.4950				0.4950
MLQLTVWGI	HIV	MN	gp160	566	A02		0.2450				0.2450
RVIEVLQRA	HIV	MN	gp160	829	A02		0.1963				0.1963
KLTPLCVTL	HIV	MN	gp160	120	A02		0.1600				0.1600
LLIAARIVEL	HIV	MN	gp160	776	A02		0.1550				0.1550
SLLNATDIAV	HIV	MN	gp160	814	A02		0.1050				0.1050
ALFLGFLGA	HIV	MN	gp160	518	A02		0.0945				0.0945
HMLQLTVWGI	HIV	MN	gp160	565	A02		0.0677				0.0677
LLNATDIAV	HIV	MN	gp160	815	A02		0.0607				0.0607
ALLYKLDIV	HIV	MN	gp160	179	A02		0.0362				0.0362
WLWYIKIFI	HIV	MN	gp160	679	A02		0.0355				0.0355
TIIVHLNESV	HIV	MN	gp160	288	A02		0.0350				0.0350
LLQYWSQEL	HIV	MN	gp160	800	A02		0.0265				0.0265
IMIVGGLVGL	HIV	MN	gp160	687	A02		0.0252				0.0252
LLYKLDIVSI	HIV	MN	gp160	180	A02		0.0245				0.0245
FLAIIWVDL	HIV	MN	gp160	753	A02		0.0233				0.0233
TLQCKIKQII	HIV	MN	gp160	415	A02		0.0200				0.0200
GLVGLRIVFA	HIV	MN	gp160	692	A02		0.0195				0.0195
FLGAAGSTM	HIV	MN	gp160	523	A02		0.0190				0.0190
IISLWDQSL	HIV	MN	gp160	107	A02		0.0179				0.0179
TVWGIKQLQA	HIV	MN	gp160	570	A02		0.0150				0.0150
LLGRRGWEV	HIV	MN	gp160	785	A02		0.0142				0.0142
AVLSIVNRV	HIV	MN	gp160	701	A02		0.0132				0.0132

Table 4

Sequence	Antigen	Strain	Molecule	Position	Motif	A1 Binding	A2 Binding	A3 Binding	A11 Binding	A24 Binding	Max. Binding
FIMIVGGLV	HIV	MN	gp160	686	A02		0.0131				0.0131
LLNATDIAVA	HIV	MN	gp160	815	A02		0.0117				0.0117
FLYGALLLA	PLP	Human		80	A02		1.9000				1.9000
SLLTFMIAA	PLP	Human		253	A02		0.5300				0.5300
FMIAATYNFAV	PLP	Human		257	A02		0.4950				0.4950
RMYGVLPWI	PLP	Human		205	A02		0.1650				0.1650
IAATYNFAV	PLP	Human		259	A02		0.0540				0.0540
GLLECCARCLV	PLP	Human		2	A02		0.0515				0.0515
YALTVVWLL	PLP	Human		157	A02		0.0415				0.0415
ALTVVWLLV	PLP	Human		158	A02		0.0390				0.0390
FLYGALLL	PLP	Human		80	A02		0.0345				0.0345
SLCADARMYGV	PLP	Human		199	A02		0.0140				0.0140
LLVFACSAV	PLP	Human		164	A02		0.0107				0.0107

45